



#4

## SEQUENCE LISTING

<110> Beetsch, Paul W.  
Avery, Angela M.  
Kaur, Balveen

<120> Broad Specificity DNA Damage Endonuclease

<130> 25-98A

<140> US/09/724,296  
<141> 2000-11-28

<150> US/09/327,984  
<151> 1999-06-08

<150> US/60/088,521  
<151> 1998-06-08

<150> US/60/134,752  
<151> 1999-05-18

<160> 71

<170> PatentIn Ver. 2.0

<210> 1  
<211> 2492  
<212> DNA  
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Coding  
sequence for fusion protein of GST signal peptide  
and the UVDE protein of *Schizosaccharomyces pombe*

<400> 1

atgacccaagt tacctatact aggttattgg aaaaattaag ggccttgc aaccactcg 60  
acttcttttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc gcatgaaagg 120  
tgataaatgg cgaaacaaaa agtttgaatt gggtttggag tttcccaatc ttccattata 180  
tattgatggt gatgttaaat taacacagtc tatgccatc atacgttata tagctgacaa 240  
gcacaacatg ttgggtgggt gtccaaaaga gcgtgcagag atttcaatgc ttgaaggagc 300  
ggttttggat attagatacg gtgttgcag aattgcatat agttaaagact ttgaaaactct 360  
caaagttgat ttcttagca agtctacgt aatgctgaaa atgttcgaag atcgttatg 420  
tcataaaaaca tatttaaatg ttgaccatgt aaccatcct gacttcatgt tgtatgacgc 480  
tcattgtatgtt gtttataca tggacccat gtgcctggat gcgttccaa aattagttt 540  
ttttaaaaaaa cgtattgaag ctatcccaca aattgataag tacttgaat ccagcaagta 600  
tatagcatgg ctttgcagg gctggcaagc cacgttgggt ggtggcgacc atcctccaaa 660  
atcgatcat ctggtccgc gtggatccat gcttaggcta ttgaaaacaa atattcaaat 720  
ctctaaacgc attgtttca ccatattaaa acaaaaaggca tttaaaggtt atcatccttg 780  
tgtaccgtcg gtttgatcca ttacttactc tcgtttcat tggttacccg atacccttaa 840

2022-06-20

aagtttactt ccaatgagct caaaaaccac actctcaatg ttaccgcaag ttaatatcg 900  
tgcgaattca ttctctgccc aaacaccagt cgacttaaaa aaagaaaatg agactgagtt 960  
agctaataatc agtgacacc tc aaaaaaaaaaag tacttctacg tctacacgaa agagggcacg 1020  
tagcagtaaa aagaaagcga cagattctgt ttccgataaa attgatgagt ctgttgcgtc 1080  
ctatgattct tcaactcatac ttaggcgatc gtcgagatca aaaaaacccgg tcaactacaa 1140  
ttccctcgta gaatccgaat cggaggagca aattagtaaa gctactaaaa aagttaaaca 1200  
aaaagaggaa gaggagtatg ttgaagaatg cgacaaaaag tctctaaaa atgaaaatg 1260  
ctctgacgag ttcgaaccgg ttgtgccgga acagttggaa actccaattt ctaaacgaag 1320  
acggtctcg tcttcgtcaa aaaatttga aaaaagatct acaatgaatc ttgatgatca 1380  
tgctccacga gagatgtttg attgtttgaa caaacccata ccctggcgag gacgattggg 1440  
gtatgctgt ttgaatacta ttttaaggc aatgaaggag agggtttt gttcacgcac 1500  
ctgcccgaatt acaaccattc aacgtgatgg gctgaaaatg gtcaagcagc tagtacgca 1560  
aaatgtttt gatttaatca aattgggtga gtggaaatcac aactttggca ttcacttcat 1620  
gagagtgagt tctgatattt ttccttcgc aagccatgca aagtatggat atacccttga 1680  
atttgcacaa tctcatctcg aggaggtggg caagctggca aataaaatata atcatcgatt 1740  
gactatgcat cctggtcagt acacccagat agcctctca cgagaagtgc tagttgattc 1800  
ggcaatacgt gatttggctt atcatgatga aattctcagt cgtatgaatg tgaatgaaca 1860  
ataaaataaa gacgctgttt taattattca ccttgggtgg acctttgaag gaaaaaaaaa 1920  
aacattggat aggttgcgt aaaaattatca acgcttgcgt gattcggta aagctcggtt 1980  
agtttttagaa aacgatgatg tttcttggc agttcaagat ttattacatt tatgccaaga 2040  
acttaatatt cctctagttt tggattggca tcatacacaac atagtgcacg gaacgcttcg 2100  
tgaaggaagt ttagatttaa tgccattaat cccaaactatt cgagaaacct ggacaagaaa 2160  
gggaattaca cagaagcaac attactcaga atcggctgat ccaacggcga tttctggat 2220  
gaaacgacgt gctcaactcg atagggtgtt tgactttcca ccgtgtgatc ctacaatgga 2280  
tctaattgata gaagotaagg aaaaggaaca ggctgttattt gaattgtgtt gacgttatg 2340  
gttacaaaat ccaccatgtc ctcttgaat tatggggctt gaatacgtc aaactcgaga 2400  
tggatattat ccgccccggag ctgaaaagcg tttaactgca agaaaaaggc gtagtagaaaa 2460  
agaagaagta gaagaggatg aaaaataaaaa at 2492

<210> 2

<211> 828

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Fusion protein  
of GST leader peptide and Schizosaccharomyces  
pombe UVDE

<400> 2

Met Thr Lys Leu Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val  
1 5 10 15

Gln Pro Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu  
20 25 30

His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe  
35 40 45

Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp  
50 55 60

Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys  
65 70 75 80

His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met  
85 90 95

Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala  
100 105 110

Tyr Ser Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu  
115 120 125

Pro Glu Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr  
130 135 140

Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala  
145 150 155 160

Leu Asp Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro  
165 170 175

Lys Leu Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp  
180 185 190

Lys Tyr Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp  
195 200 205

Gln Ala Thr Phe Gly Gly Asp His Pro Pro Lys Ser Asp His Leu  
210 215 220

Val Pro Arg Gly Ser Met Leu Arg Leu Leu Lys Arg Asn Ile Gln Ile  
225 230 235 240

Ser Lys Arg Ile Val Phe Thr Ile Leu Lys Gln Lys Ala Phe Lys Gly  
245 250 255

Asn His Pro Cys Val Pro Ser Val Cys Thr Ile Thr Tyr Ser Arg Phe  
260 265 270

His Cys Leu Pro Asp Thr Leu Lys Ser Leu Leu Pro Met Ser Ser Lys  
275 280 285

Thr Thr Leu Ser Met Leu Pro Gln Val Asn Ile Gly Ala Asn Ser Phe  
290 295 300

Ser Ala Glu Thr Pro Val Asp Leu Lys Lys Glu Asn Glu Thr Glu Leu  
305 310 315 320

Ala Asn Ile Ser Gly Pro His Lys Lys Ser Thr Ser Thr Ser Thr Arg  
325 330 335

Lys Arg Ala Arg Ser Ser Lys Lys Ala Thr Asp Ser Val Ser Asp  
340 345 350

Lys Ile Asp Glu Ser Val Ala Ser Tyr Asp Ser Ser Thr His Leu Arg  
355 360 365

Arg Ser Ser Arg Ser Lys Lys Pro Val Asn Tyr Asn Ser Ser Ser Glu  
370 375 380

Ser Glu Ser Glu Glu Gln Ile Ser Lys Ala Thr Lys Lys Val Lys Gln  
385 390 395 400

Lys Glu Glu Glu Glu Tyr Val Glu Glu Val Asp Glu Lys Ser Leu Lys  
405 410 415

Asn Glu Ser Ser Ser Asp Glu Phe Glu Pro Val Val Pro Glu Gln Leu  
420 425 430

Glu Thr Pro Ile Ser Lys Arg Arg Arg Ser Arg Ser Ser Ala Lys Asn  
435 440 445

Leu Glu Lys Glu Ser Thr Met Asn Leu Asp Asp His Ala Pro Arg Glu  
450 455 460

Met Phe Asp Cys Leu Asp Lys Pro Ile Pro Trp Arg Gly Arg Leu Gly  
465 470 475 480

Tyr Ala Cys Leu Asn Thr Ile Leu Arg Ser Met Lys Glu Arg Val Phe  
485 490 495

Cys Ser Arg Thr Cys Arg Ile Thr Thr Ile Gln Arg Asp Gly Leu Glu  
500 505 510

Ser Val Lys Gln Leu Gly Thr Gln Asn Val Leu Asp Leu Ile Lys Leu  
515 520 525

Val Glu Trp Asn His Asn Phe Gly Ile His Phe Met Arg Val Ser Ser  
530 535 540

Asp Leu Phe Pro Phe Ala Ser His Ala Lys Tyr Gly Tyr Thr Leu Glu  
545 550 555 560

Phe Ala Gln Ser His Leu Glu Glu Val Gly Lys Leu Ala Asn Lys Tyr  
565 570 575

Asn His Arg Leu Thr Met His Pro Gly Gln Tyr Thr Gln Ile Ala Ser  
580 585 590

Pro Arg Glu Val Val Val Asp Ser Ala Ile Arg Asp Leu Ala Tyr His  
595 600 605

Asp Glu Ile Leu Ser Arg Met Lys Leu Asn Glu Gln Leu Asn Lys Asp  
610 615 620

Ala Val Leu Ile Ile His Leu Gly Gly Thr Phe Glu Gly Lys Lys Glu  
625 630 635 640

Thr Leu Asp Arg Phe Arg Lys Asn Tyr Gln Arg Leu Ser Asp Ser Val  
645 650 655

Lys Ala Arg Leu Val Leu Glu Asn Asp Asp Val Ser Trp Ser Val Gln  
660 665 670

Asp Leu Leu Pro Leu Cys Gln Glu Leu Asn Ile Pro Leu Val Leu Asp  
675 680 685

Trp His His His Asn Ile Val Pro Gly Thr Leu Arg Glu Gly Ser Leu  
690 695 700

Asp Leu Met Pro Leu Ile Pro Thr Ile Arg Glu Thr Trp Thr Arg Lys  
705 710 715 720

Gly Ile Thr Gln Lys Gln His Tyr Ser Glu Ser Ala Asp Pro Thr Ala  
725 730 735

Ile Ser Gly Met Lys Arg Arg Ala His Ser Asp Arg Val Phe Asp Phe  
740 745 750

Pro Pro Cys Asp Pro Thr Met Asp Leu Met Ile Glu Ala Lys Glu Lys  
755 760 765

Glu Gln Ala Val Phe Glu Leu Cys Arg Arg Tyr Glu Leu Gln Asn Pro  
770 775 780

Pro Cys Pro Leu Glu Ile Met Gly Pro Glu Tyr Asp Gln Thr Arg Asp  
785 790 795 800

Gly Tyr Tyr Pro Pro Gly Ala Glu Lys Arg Leu Thr Ala Arg Lys Arg  
805 810 815

Arg Ser Arg Lys Glu Glu Val Glu Asp Glu Lys  
820 825

<210> 3

<211> 1161

<212> DNA

<213> Schizosaccharomyces pombe

<220>

<221> misc\_feature

<222> (1)..(1161)

<223> DNA sequence encoding UVDE protein, truncated at amino acid residue 228.

<400> 3

gatgatcatg ctccacgaga gatgttgat tggggaca aaccatacc ctggcgagg 60  
cgattgggt atgcttgtt gaatactatt ttaaggtcaa tgaaggagag gttttttgt 120  
tcacgcacct gccgaattac aaccattcaa cgtgatgggc tcgaaagtgt caagcagcta 180  
ggtacgaaa atgtttaga ttaatcaaa ttgggtgagt ggaatcaca ctttggcatt 240  
cacttcatga gagtgagttc tgatttattt ctttcgcaa gccatcaca gtatggatat 300  
acccttgaat ttgcacaatc tcatctcgag gaggtgggca agctggcaa taaatataat 360  
catcgattga ctatgcattc tggtcagttt acccagatag cctctccacg agaagtcgta 420  
gttggattcgg caatacgtga ttggcttat catgatgaaa ttctcagtcg tatgaagttg 480  
aatgaacaat taaataaaga cgctgttttta attattcacc ttgggtgtac ctttgaagga 540  
aaaaaaagaaa cattggatag gtttcgtaaa aattatcaac gttgtctga ttcggtaaa 600  
gctcggttag ttttagaaaa cgtatgtt tcttggtcag ttcaagattt attaccttta 660  
tgccaagaac ttaatattcc tctagtttg gattggcattt atcacaacat agtgcagg 720  
acgcttcgtg aaggaagttt agatttaatg coattaatcc caactattcg agaaacctgg 780  
acaagaaaagg gaattacaca gaagcaacat tactcagaat cggctgatcc aacggcgatt 840  
tctggatga aacgacgtgc tcactctgtt aggggtttt actttccacc gtgtgatcct 900  
acaatggatc taatgtataga agctaaggaa aaggaacagg ctgtatgttga attgtgtaga 960  
cgttatgtagt tacaaaatcc accatgtcctt cttgaaatattt tggggcctga atacgatcaa 1020  
actcggatgt gatattatcc gcccggagct gaaaagcggt taactgcaag aaaaaggcgt 1080  
agtagaaaaag aagaagttaga agaggatgaa aaataaaaaat ccgtcataact ttttgattta 1140  
tggcataatt tagccatctc c 1161

<210> 4

<211> 371

<212> PRT

<213> *Schizosaccharomyces pombe*

<220>

<221> VARIANT

<222> (1) . . (371)

<223> Truncated version of the UVDE protein.

<400> 4

Asp Asp His Ala Pro Arg Glu Met Phe Asp Cys Leu Asp Lys Pro Ile  
1 5 10 15

Pro Trp Arg Gly Arg Leu Gly Tyr Ala Cys Leu Asn Thr Ile Leu Arg  
20 25 30

Ser Met Lys Glu Arg Val Phe Cys Ser Arg Thr Cys Arg Ile Thr Thr  
           35                  40                  45

Ile Gln Arg Asp Gly Leu Glu Ser Val Lys Gln Leu Gly Thr Gln Asn  
50 55 60

Val Leu Asp Leu Ile Lys Leu Val Glu Trp Asn His Asn Phe Gly Ile  
65 70 75 80

His Phe Met Arg Val Ser Ser Asp Leu Phe Pro Phe Ala Ser His Ala  
85 90 95

Lys Tyr Gly Tyr Thr Leu Glu Phe Ala Gln Ser His Leu Glu Glu Val  
100 105 110

Gly Lys Leu Ala Asn Lys Tyr Asn His Arg Leu Thr Met His Pro Gly  
115 120 125

Gln Tyr Thr Gln Ile Ala Ser Pro Arg Glu Val Val Val Asp Ser Ala  
130 135 140

Ile Arg Asp Leu Ala Tyr His Asp Glu Ile Leu Ser Arg Met Lys Leu  
145 150 155 160

Asn Glu Gln Leu Asn Lys Asp Ala Val Leu Ile Ile His Leu Gly Gly  
165 170 175

Thr Phe Glu Gly Lys Lys Glu Thr Leu Asp Arg Phe Arg Lys Asn Tyr  
180 185 190

Gln Arg Leu Ser Asp Ser Val Lys Ala Arg Leu Val Leu Glu Asn Asp  
195 200 205

Asp Val Ser Trp Ser Val Gln Asp Leu Leu Pro Leu Cys Gln Glu Leu  
210 215 220

Asn Ile Pro Leu Val Leu Asp Trp His His His Asn Ile Val Pro Gly  
225 230 235 240

Thr Leu Arg Glu Gly Ser Leu Asp Leu Met Pro Leu Ile Pro Thr Ile  
245 250 255

Arg Glu Thr Trp Thr Arg Lys Gly Ile Thr Gln Lys Gln His Tyr Ser  
260 265 270

Glu Ser Ala Asp Pro Thr Ala Ile Ser Gly Met Lys Arg Arg Ala His  
275 280 285

Ser Asp Arg Val Phe Asp Phe Pro Pro Cys Asp Pro Thr Met Asp Leu  
290 295 300

Met Ile Glu Ala Lys Glu Lys Glu Gln Ala Val Phe Glu Leu Cys Arg  
305 310 315 320

Arg Tyr Glu Leu Gln Asn Pro Pro Cys Pro Leu Glu Ile Met Gly Pro  
325 330 335

Glu Tyr Asp Gln Thr Arg Asp Gly Tyr Tyr Pro Pro Gly Ala Glu Lys  
340 345 350

Arg Leu Thr Ala Arg Lys Arg Arg Ser Arg Lys Glu Glu Val Glu Glu  
355 360 365

Asp Glu Lys  
370

```
<210> 5
<211> 1811
<212> DNA
<213> Artificial Sequence
```

<220>  
<223> Description of Artificial Sequence: Nucleotide  
sequence encoding fusion protein of GST signal  
peptide and the truncated version of *S. pombe* UVDE  
protein.

<400> 5  
atgaccaagt tacctatact aggttattgg aaaaattaag ggccttgc aaccacactcg 60  
acttctttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc gcgatgaagg 120  
tgataaatgg cggaaacaaaa agtttgaatt gggtttggag tttcccaatc ttccttatta 180  
tattgtatggt gatgttaaat taacacagtc tatggccatc atacgttata tagctgacaa 240  
gcacaacatg ttgggttgggt gtccaaaaga gcgtgcagag atttcaatgc ttgaaggagc 300  
ggttttggat attagatacg gtgttcgag aattgcataat agtaaagact ttgaaactct 360  
caaagttgat ttcttagca agctacactga aatgctgaaa atgttcgaag atcggttatg 420  
tcataaaaaca tatttaaatg ttgaccatgt aaccatctt gacttcatgt tttatgacgc 480  
tcttgatgtt gtttataca tggaccaat gtgcctggat gcgttccaa aatttagttg 540  
ttttaaaaaa cgatttgaag ctatcccaca aattgataag tacttgaat ccagcaagta 600  
tatagcatgg cctttgcagg gctggcaagc cacgttggat ggtggcgacc atccctccaa 660  
atcgatcatat ctgggttccgc gtggatccga tggatcatgtt ccacgagaga ttttggattg 720  
tttggacaaa cccataccct ggcgaggacg attggggat gcttggatgataactt 780  
aaggtaatg aaggagaggg tttttgttc acgcacactgc cgaattacaa ccattcaacg 840  
tgatgggctc gaaaagtgtca agcagctagg tacgaaaaat gttttagatt taatcaaatt 900  
ggttgagtgg aatcacaact ttggcattca cttcatgaga gtgagttctg atttatttcc 960  
tttcgcaagc catgcaaatg atggatatac ctttgcattt gcacaatctc atctcgagga 1020  
ggtgggcaag ctggcaataa aatataatca tcgattgact atgcacatctg gtcagttacac 1080  
ccagatagcc tctccacgag aagtgttagt tgattcggca atacgtgatt tggcttatca 1140  
tgatggaaatt ctcaatgttca tgaagttgaa tgaacaattt aataaagacg ctgttttaat 1200  
tattcacctt ggtggatctt ttgaaggaaa aaaagaaaaca ttggataggt ttctgtaaaa 1260  
ttatcaacgc ttgtctgatt cggtaaagc tcgttttagtt ttagaaaacg atgatgtttc 1320  
ttggtcagtt caagatttat tacctttatg ccaagaactt aatattccct tagtttgg 1380  
ttggcatcat cacaacatag tgccaggaac gttcgtgaa ggaagtttag atttaatgcc 1440  
attaatccca actattcgag aaacctggac aagaaaaggaa attacacaga agcaacatta 1500  
ctcagaatcg gctgatccaa cggcgatttc tggatggaaa cgacgtgctc actctgtatag 1560  
ggtgtttgac ttccacccgt gtgatcctac aatggatcta atgatagaag ctaaggaaaa 1620  
ggaacaggct gtatgttcaat tggatgtacg ttatgtatggataaaatccac catgtccct 1680  
tgaattatg gggcctgaat acgatcaaac tcgagatgga tattatccgc ccggagctga 1740  
aaagcgatccatc actgcaagaa aaaggcgtag tagaaaagaa gaagtagaaaggatgaaaa 1800  
ataaggatccatc 1811

<210> 6  
<211> 600  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Fusion protein comprising the GST signal peptide and the truncated UVDE protein of *S. pombe*.

<400> 6  
Met Thr Lys Leu Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val  
1 5 10 15  
Gln Pro Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu  
20 25 30  
His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe  
35 40 45  
Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp  
50 55 60  
Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys  
65 70 75 80  
His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met  
85 90 95  
Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala  
100 105 110  
Tyr Ser Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu  
115 120 125  
Pro Glu Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr  
130 135 140  
Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala  
145 150 155 160  
Leu Asp Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro  
165 170 175  
Lys Leu Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp  
180 185 190  
Lys Tyr Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp  
195 200 205

Gln Ala Thr Phe Gly Gly Asp His Pro Pro Lys Ser Asp His Leu  
210 215 220

Val Pro Arg Gly Ser Asp Asp His Ala Pro Arg Glu Met Phe Asp Cys  
225 230 235 240

Leu Asp Lys Pro Ile Pro Trp Arg Gly Arg Leu Gly Tyr Ala Cys Leu  
245 250 255

Asn Thr Ile Leu Arg Ser Met Lys Glu Arg Val Phe Cys Ser Arg Thr  
260 265 270

Cys Arg Ile Thr Thr Ile Gln Arg Asp Gly Leu Glu Ser Val Lys Gln  
275 280 285

Leu Gly Thr Gln Asn Val Leu Asp Leu Ile Lys Leu Val Glu Trp Asn  
290 295 300

His Asn Phe Gly Ile His Phe Met Arg Val Ser Ser Asp Leu Phe Pro  
305 310 315 320

Phe Ala Ser His Ala Lys Tyr Gly Tyr Thr Leu Glu Phe Ala Gln Ser  
325 330 335

His Leu Glu Glu Val Gly Lys Leu Ala Asn Lys Tyr Asn His Arg Leu  
340 345 350

Thr Met His Pro Gly Gln Tyr Thr Gln Ile Ala Ser Pro Arg Glu Val  
355 360 365

Val Val Asp Ser Ala Ile Arg Asp Leu Ala Tyr His Asp Glu Ile Leu  
370 375 380

Ser Arg Met Lys Leu Asn Glu Gln Leu Asn Lys Asp Ala Val Leu Ile  
385 390 395 400

Ile His Leu Gly Gly Thr Phe Glu Gly Lys Lys Glu Thr Leu Asp Arg  
405 410 415

Phe Arg Lys Asn Tyr Gln Arg Leu Ser Asp Ser Val Lys Ala Arg Leu  
420 425 430

Val Leu Glu Asn Asp Asp Val Ser Trp Ser Val Gln Asp Leu Leu Pro  
435 440 445

Leu Cys Gln Glu Leu Asn Ile Pro Leu Val Leu Asp Trp His His His  
450 455 460

Asn Ile Val Pro Gly Thr Leu Arg Glu Gly Ser Leu Asp Leu Met Pro  
465 470 475 480

Leu Ile Pro Thr Ile Arg Glu Thr Trp Thr Arg Lys Gly Ile Thr Gln  
485 490 495

Lys Gln His Tyr Ser Glu Ser Ala Asp Pro Thr Ala Ile Ser Gly Met  
500 505 510

Lys Arg Arg Ala His Ser Asp Arg Val Phe Asp Phe Pro Pro Cys Asp  
515 520 525

Pro Thr Met Asp Leu Met Ile Glu Ala Lys Glu Lys Glu Gln Ala Val  
530 535 540

Phe Glu Leu Cys Arg Arg Tyr Glu Leu Gln Asn Pro Pro Cys Pro Leu  
545 550 555 560

Glu Ile Met Gly Pro Glu Tyr Asp Gln Thr Arg Asp Gly Tyr Tyr Pro  
565 570 575

Pro Gly Ala Glu Lys Arg Leu Thr Ala Arg Lys Arg Arg Ser Arg Lys  
580 585 590

Glu Glu Val Glu Glu Asp Glu Lys  
595 600

<210> 7

<211> 688

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Nucleotide  
sequence encoding GST signal peptide.

<400> 7

atgaccaagt tacctatact aggttattgg aaaaattaag ggccttgc aaccactcg 60  
acttctttg gaatatcttg aagaaaaata tgaagagcat ttgtatgagc gcgatgaagg 120  
tgataaatgg cgaacaaaaa agttgaatt gggtttggag ttcccaatc ttccttatta 180  
tattgatggt gatgttaaat taacacagtc tatggccatc atacgtata tagctgacaa 240  
gcacaacatg ttgggtgggt gtccaaaaga gcgtgcagag atttcaatgc ttgaaggagc 300  
gttttggat attagatacg gtgttgcag aattgcatat agtaaagact ttgaaactct 360  
caaagttgat tttcttagca agctacctga aatgctgaaa atgttgcag atcgtttatg 420  
tcataaaaaca tatttaaatg ttgaccatgt aaccatcct gacttcatgt tgtatgacgc 480  
tcttgatggt gtttataca tggacccaaat gtgcctggat gcgttcccaa aattagttg 540  
ttttaaaaaa cgtattgaag ctatccaca aattgataag tacttgaat ccagcaagta 600  
tatagcatgg ccttgcagg gctggcaagc cacgttggc ggtggcgacc atcctccaaa 660  
atcgatcat ctggttccgc gtggatcc 688

<210> 8

<211> 229

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: Amino acid sequence of GST signal peptide

<400> 8

Met Thr Lys Leu Pro Ile Leu Gly Tyr Trp Lys Ile Lys Gly Leu Val  
1 5 10 15

Gln Pro Thr Arg Leu Leu Leu Glu Tyr Leu Glu Glu Lys Tyr Glu Glu  
20 25 30

His Leu Tyr Glu Arg Asp Glu Gly Asp Lys Trp Arg Asn Lys Lys Phe  
35 40 45

Glu Leu Gly Leu Glu Phe Pro Asn Leu Pro Tyr Tyr Ile Asp Gly Asp  
50 55 60

Val Lys Leu Thr Gln Ser Met Ala Ile Ile Arg Tyr Ile Ala Asp Lys  
65 70 75 80

His Asn Met Leu Gly Gly Cys Pro Lys Glu Arg Ala Glu Ile Ser Met  
85 90 95

Leu Glu Gly Ala Val Leu Asp Ile Arg Tyr Gly Val Ser Arg Ile Ala  
100 105 110

Tyr Ser Lys Asp Phe Glu Thr Leu Lys Val Asp Phe Leu Ser Lys Leu  
115 120 125

Pro Glu Met Leu Lys Met Phe Glu Asp Arg Leu Cys His Lys Thr Tyr  
130 135 140

Leu Asn Gly Asp His Val Thr His Pro Asp Phe Met Leu Tyr Asp Ala  
145 150 155 160

Leu Asp Val Val Leu Tyr Met Asp Pro Met Cys Leu Asp Ala Phe Pro  
165 170 175

Lys Leu Val Cys Phe Lys Lys Arg Ile Glu Ala Ile Pro Gln Ile Asp  
180 185 190

Lys Tyr Leu Lys Ser Ser Lys Tyr Ile Ala Trp Pro Leu Gln Gly Trp  
195 200 205

Gln Ala Thr Phe Gly Gly Asp His Pro Pro Lys Ser Asp His Leu  
210 215 220

Val Pro Arg Gly Ser  
225

<210> 9  
<211> 36  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 9  
tgaggatcca atcgaaaaa tttttttttt cttagg 36

<210> 10  
<211> 23  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 10  
ggccatggtt atttttcatc ctc 23

<210> 11  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 11  
aatgggatcc gatgatcatg ctccacga 28

<210> 12  
<211> 28  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 12  
gggatccatcc tttttcatcc tcttctac 28

<210> 13  
<211> 30

102300 9622260

<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing cis-syn  
cyclobutane pyrimidine dimer.

<220>  
<221> misc\_feature  
<222> (15)..(16)  
<223> At positions 15- 16, the T-T is in the form of  
cis-syn cyclobutane pyrimidine dimer.

<400> 13  
catgcctgca cgaatatacg aattcgtaat 30

<210> 14  
<211> 30  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Undamaged  
double stranded oligonucleotide.

<400> 14  
catgcctgca cgaatatacg aattcgtaat 30

<210> 15  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing cis-syn  
cyclobutane dimer at positions 21-22.

<400> 15  
agctaccatg cctgcacgaa ttaagcaatt cgtaatcatg gtcatacg 49

<210> 16  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing cis-syn  
cyclobutane pyrimidine dimer at positions 21-22.

092246 05201

<400> 16  
agctaccatg cctgcacgaa ttaagcaatt cgtaatcatg gtcatagct 49  
  
<210> 17  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing trans- syn 11  
cyclobutane pyrimidine dimer at positions 21-22.  
  
<400> 17  
agctaccatg cctgcacgaa ttaagcaatt cgtaatcatg gtcatagct 49  
  
<210> 18  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing a 6-4 photo  
product at position 21-22.  
  
<400> 18  
agctaccatg cctgcacgaa ttaagcaatt cgtaatcatg gtcatagct 49  
  
<210> 19  
<211> 49  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing a Dewar  
isomer.  
  
<400> 19  
agctaccatg cctgcacgaa ttaagcaatt cgtaatcatg gtcatagct 49  
  
<210> 20  
<211> 32  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide containing cisplatin DNA  
diadduct at position 16-17.

<400> 20  
tccctccttc cttccggccc tccttccct tc 32

<210> 21  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n is uracil at  
position 21.

<220>  
<221> misc\_feature  
<222> (21)  
<223> The n at position 21 is uracil.

<400> 21  
cttggactgg atgtcggcac nacggatac aggagca 37

<210> 22  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n is  
dihydrouracil at position 21.

<220>  
<221> misc\_feature  
<222> (21)  
<223> At position 21, n is dihydrouracil.

<400> 22  
cttggactgg atgtcggcac nacggatac aggagca 37

<210> 23  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 21  
represents an abasic site.

<220>  
<221> misc\_feature

<222> (21)  
<223> At position 21, n is an abasic site.

<400> 23  
cttggactgg atgtcggcac nacggatac aggagca 37

<210> 24  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 13  
is an inosine.

<220>  
<221> misc\_feature  
<222> (13)  
<223> At position 13, n is inosine.

<400> 24  
tgcaggtcga ctnaggagga tccccgggta c 31

<210> 25  
<211> 31  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 13  
is xanthine.

<220>  
<221> misc\_feature  
<222> (13)  
<223> N at position 13 is xanthine.

<400> 25  
tgcaggtcga ctnaggagga tccccgggta c 31

<210> 26  
<211> 37  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 21  
is 8-oxoguanine.

```
<220>
<221> misc_feature
<222> (21)
<223> N at position 21 is 8-oxoguanine.

<400> 26
cttggactgg atgtcggcac nagcggatac aggagca

<210> 27
<211> 30
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Double
      stranded oligonucleotide representing all 16
      possible base pair mismatches at position 18 in
      individual preparations.

<220>
<221> misc_feature
<222> (18)
<223> N at position 18 represents all 16 possible base
      pair mismatches.

<400> 27
gtacccgggg atcctccnag tcgacctgca

<210> 28
<211> 41
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Double
      stranded oligonucleotide containing a CA
      mismatched base pair at position 21.

<220>
<221> misc_feature
<222> (21)
<223> N at position 21 represents C of C/A mismatched
      base pair.

<400> 28
cgtagcatg cctgcacgaa ntaagcaatt cgtaatgcat t

<210> 29
<211> 41
<212> DNA
<213> Artificial Sequence
```

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein there is a C/A  
mismatched base pair at position 36.

<220>  
<221> misc\_feature  
<222> (36)  
<223> N at position 36 represents a C/A mismatched base  
pair.

<400> 29  
cgttacaagt ccgtcacgaa ttaagcaatt cgtaangcat t

41

<210> 30  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 31  
represents a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (31)  
<223> The n at position 31 represents C of C/A  
mismatched base pair.

<400> 30  
cgttacaagt ccgtcacgaa ttaagcaatt ngtaacgcat t

41

<210> 31  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 26  
is a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (26)  
<223> N at position 26 represents a C/A mismatched base  
pair.

<400> 31  
cgttacaagt ccgtcacgaa ttaagnaatt cgtaacgcat t

41

03/24/96 10:25 AM  
TO:  
  
<210> 32  
<211> 41  
<212> DNA  
<213> Artificial Sequence  
  
<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 21  
is a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (21)  
<223> The n at position 21 represents a C/A mismatched  
base pair.

<400> 32  
cgttacaagt ccgtcacgac ntaagcaatt cgtaacgcat t

41

<210> 33  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 15  
represents a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (15)  
<223> The n at position 15 represents a C/A mismatched  
base pair.

<400> 33  
cgttacaagt ccgttnacgaa ttaagcaatt cgtaacgcat t

41

<210> 34  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double  
stranded oligonucleotide wherein n at position 10  
is a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (10)

<223> The n at position 10 represents a C/A mismatched base pair.

<400> 34  
cgttacaagn ccgtcacgaa ttaagcaatt cgtaacgcatt

41

<210> 35  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence: Double stranded oligonucleotide wherein n at position 5 is a C/A mismatched base pair.

<220>  
<221> misc\_feature  
<222> (5)  
<223> The n at position 5 represents a C/A mismatched base pair.

<400> 35  
cgtncaagt ccgtcacgaa ttaagcaatt cgtaacgcatt

41

<210> 36  
<211> 656  
<212> PRT  
<213> Neurospora crassa

<400> 36  
Met Pro Ser Arg Lys Ser Lys Ala Ala Ala Leu Asp Thr Pro Gln Ser  
1 5 10 15

Glu Ser Ser Thr Phe Ser Ser Thr Leu Asp Ser Ser Ala Pro Ser Pro  
20 25 30

Ala Arg Asn Leu Arg Arg Ser Gly Arg Asn Ile Leu Gln Pro Ser Ser  
35 40 45

Glu Lys Asp Arg Asp His Glu Lys Arg Ser Gly Glu Glu Leu Ala Gly  
50 55 60

Arg Met Met Gly Lys Asp Ala Asn Gly His Cys Leu Arg Glu Gly Lys  
65 70 75 80

Glu Gln Glu Glu Gly Val Lys Met Ala Ile Glu Gly Leu Ala Arg Met  
85 90 95

Glu Arg Arg Leu Gln Arg Ala Thr Lys Arg Gln Lys Lys Gln Leu Glu  
100 105 110

Glu Asp Gly Ile Pro Val Pro Ser Val Val Ser Arg Phe Pro Thr Ala  
115 120 125  
Pro Tyr His His Lys Ser Thr Asn Ala Glu Glu Arg Glu Ala Lys Glu  
130 135 140  
Pro Val Leu Lys Thr His Ser Lys Asp Val Glu Arg Glu Ala Glu Ile  
145 150 155 160  
Gly Val Asp Asp Val Val Lys Met Glu Pro Ala Ala Thr Asn Ile Ile  
165 170 175  
Glu Pro Glu Asp Ala Gln Asp Ala Ala Glu Arg Gly Ala Ala Arg Pro  
180 185 190  
Pro Ala Val Asn Ser Ser Tyr Leu Pro Leu Pro Trp Lys Gly Arg Leu  
195 200 205  
Gly Tyr Ala Cys Leu Asn Thr Tyr Leu Arg Asn Ala Lys Pro Pro Ile  
210 215 220  
Phe Ser Ser Arg Thr Cys Arg Met Ala Ser Ile Val Asp His Arg His  
225 230 235 240  
Pro Leu Gln Phe Glu Asp Glu Pro Glu His His Leu Lys Asn Lys Pro  
245 250 255  
Asp Lys Ser Lys Glu Pro Gln Asp Glu Leu Gly His Lys Phe Val Gln  
260 265 270  
Glu Leu Gly Leu Ala Asn Ala Arg Asp Ile Val Lys Met Leu Cys Trp  
275 280 285  
Asn Glu Lys Tyr Gly Ile Arg Phe Leu Arg Leu Ser Ser Glu Met Phe  
290 295 300  
Pro Phe Ala Ser His Pro Val His Gly Tyr Lys Leu Ala Pro Phe Ala  
305 310 315 320  
Ser Glu Val Leu Ala Glu Ala Gly Arg Val Ala Ala Glu Leu Gly His  
325 330 335  
Arg Leu Thr Thr His Pro Gly Gln Phe Thr Gln Leu Gly Ser Pro Arg  
340 345 350  
Lys Glu Val Val Glu Ser Ala Ile Arg Asp Leu Glu Tyr His Asp Glu  
355 360 365  
Leu Leu Ser Leu Leu Lys Leu Pro Glu Gln Gln Asn Arg Asp Ala Val  
370 375 380

Met Ile Ile His Met Gly Gly Gln Phe Gly Asp Lys Ala Ala Thr Leu  
385 390 395 400

Glu Arg Phe Lys Arg Asn Tyr Ala Arg Leu Ser Gln Ser Cys Lys Asn  
405 410 415

Arg Leu Val Leu Glu Asn Asp Asp Val Gly Trp Thr Val His Asp Leu  
420 425 430

Leu Pro Val Cys Glu Glu Leu Asn Ile Pro Met Val Leu Asp Tyr His  
435 440 445

His His Asn Ile Cys Phe Asp Pro Ala His Leu Arg Glu Gly Thr Leu  
450 455 460

Asp Ile Ser Asp Pro Lys Leu Gln Glu Arg Ile Ala Asn Thr Trp Lys  
465 470 475 480

Arg Lys Gly Ile Lys Gln Lys Met His Tyr Ser Glu Pro Cys Asp Gly  
485 490 495

Ala Val Thr Pro Arg Asp Arg Arg Lys His Arg Pro Arg Val Met Thr  
500 505 510

Leu Pro Pro Cys Pro Pro Asp Met Asp Leu Met Ile Glu Ala Lys Asp  
515 520 525

Lys Glu Gln Ala Val Phe Glu Leu Met Arg Thr Phe Lys Leu Pro Gly  
530 535 540

Phe Glu Lys Ile Asn Asp Met Val Pro Tyr Asp Arg Asp Asp Glu Asn  
545 550 555 560

Arg Pro Ala Pro Pro Val Lys Ala Pro Lys Lys Lys Gly Gly Lys  
565 570 575

Arg Lys Arg Thr Thr Asp Glu Glu Ala Ala Glu Pro Glu Glu Val Asp  
580 585 590

Thr Ala Ala Asp Asp Val Lys Asp Ala Pro Glu Gly Pro Lys Glu Val  
595 600 605

Pro Glu Glu Glu Arg Ala Met Gly Gly Pro Tyr Asn Arg Val Tyr Trp  
610 615 620

Pro Leu Gly Cys Glu Glu Trp Leu Lys Pro Lys Lys Arg Glu Val Lys  
625 630 635 640

Lys Gly Lys Val Pro Glu Glu Val Glu Asp Glu Gly Glu Phe Asp Gly  
645 650 655

<210> 37  
<211> 317  
<212> PRT  
<213> *Bacillus subtilis*

<400> 37  
 Met Ile Phe Arg Phe Gly Phe Val Ser Asn Ala Met Ser Leu Trp Asp  
 1 5 10 15  
 Ala Ser Pro Ala Lys Thr Leu Thr Phe Ala Arg Tyr Ser Lys Leu Ser  
 20 25 30  
 Lys Thr Glu Arg Lys Glu Ala Leu Leu Thr Val Thr Lys Ala Asn Leu  
 35 40 45  
 Arg Asn Thr Met Arg Thr Leu His Tyr Ile Ile Gly His Gly Ile Pro  
 50 55 60  
 Leu Tyr Arg Phe Ser Ser Ile Val Pro Leu Ala Thr His Pro Asp  
 65 70 75 80  
 Val Met Trp Asp Phe Val Thr Pro Phe Gln Lys Glu Phe Arg Glu Ile  
 85 90 95  
 Gly Glu Leu Val Lys Thr His Gln Leu Arg Thr Ser Phe His Pro Asn  
 100 105 110  
 Gln Phe Thr Leu Phe Thr Ser Pro Lys Glu Ser Val Thr Lys Asn Ala  
 115 120 125  
 Val Thr Asp Met Ala Tyr His Tyr Arg Met Leu Glu Ala Met Gly Ile  
 130 135 140  
 Ala Asp Arg Ser Val Ile Asn Ile His Ile Gly Gly Ala Tyr Gly Asn  
 145 150 155 160  
 Lys Asp Thr Ala Thr Ala Gln Phe His Gln Asn Ile Lys Gln Leu Pro  
 165 170 175  
 Gln Glu Ile Lys Glu Arg Met Thr Leu Glu Asn Asp Asp Lys Thr Tyr  
 180 185 190  
 Thr Thr Glu Glu Thr Leu Gln Val Cys Glu Gln Glu Asp Val Pro Phe  
 195 200 205  
 Val Phe Asp Phe His His Phe Tyr Ala Asn Pro Asp Asp His Ala Asp  
 210 215 220

Leu Asn Val Ala Leu Pro Arg Met Ile Lys Thr Trp Glu Arg Ile Gly  
225 230 235 240

Leu Gln Pro Lys Val His Leu Ser Ser Pro Lys Ser Glu Gln Ala Ile  
245 250 255

Arg Ser His Ala Asp Tyr Val Asp Ala Asn Phe Leu Leu Glu Arg Phe  
260 265 270

Arg Gln Trp Gly Thr Asn Ile Asp Phe Met Ile Glu Ala Lys Gln Lys  
275 280 285

Asp Lys Ala Leu Leu Arg Leu Met Asp Glu Leu Ser Ser Ile Arg Gly  
290 295 300

Val Lys Arg Ile Gly Gly Ala Leu Gln Trp Lys Ser  
305 310 315

<210> 38

<211> 580

<212> PRT

<213> Homo sapiens

<400> 38  
Met Gly Thr Thr Gly Leu Glu Ser Leu Ser Leu Gly Asp Arg Gly Ala  
1 5 10 15

Ala Pro Thr Val Thr Ser Ser Glu Arg Leu Val Pro Asp Pro Pro Asn  
20 25 30

Asp Leu Arg Lys Glu Asp Val Ala Met Glu Leu Glu Arg Val Gly Glu  
35 40 45

Asp Glu Glu Gln Met Met Ile Lys Arg Ser Ser Glu Cys Asn Pro Leu  
50 55 60

Leu Gln Glu Pro Ile Ala Ser Ala Gln Phe Gly Ala Thr Ala Gly Thr  
65 70 75 80

Glu Cys Arg Lys Ser Val Pro Cys Gly Trp Glu Arg Val Val Lys Gln  
85 90 95

Arg Leu Phe Gly Lys Thr Ala Gly Arg Phe Asp Val Tyr Phe Ile Ser  
100 105 110

Pro Gln Gly Leu Lys Phe Arg Ser Lys Ser Ser Leu Ala Asn Tyr Leu  
115 120 125

His Lys Asn Gly Glu Thr Ser Leu Lys Pro Glu Asp Phe Asp Phe Thr  
130 135 140

Val Leu Ser Lys Arg Gly Ile Lys Ser Arg Tyr Lys Asp Cys Ser Met  
145 150 155 160

Ala Ala Leu Thr Ser His Leu Gln Asn Gln Ser Asn Asn Ser Asn Trp  
165 170 175

Asn Leu Arg Thr Arg Ser Lys Cys Lys Lys Asp Val Phe Met Pro Pro  
180 185 190

Ser Ser Ser Glu Leu Gln Glu Ser Arg Gly Leu Ser Asn Phe Thr  
195 200 205

Ser Thr His Leu Leu Leu Lys Glu Asp Glu Gly Val Asp Asp Val Asn  
210 215 220

Phe Arg Lys Val Arg Lys Pro Lys Gly Lys Val Thr Ile Leu Lys Gly  
225 230 235 240

Ile Pro Ile Lys Lys Thr Lys Lys Gly Cys Arg Lys Ser Cys Ser Gly  
245 250 255

Phe Val Gln Ser Asp Ser Lys Arg Glu Ser Val Cys Asn Lys Ala Asp  
260 265 270

Ala Glu Ser Glu Pro Val Ala Gln Lys Ser Gln Leu Asp Arg Thr Val  
275 280 285

Cys Ile Ser Asp Ala Gly Ala Cys Gly Glu Thr Leu Ser Val Thr Ser  
290 295 300

Glu Glu Asn Ser Leu Val Lys Lys Glu Arg Ser Leu Ser Ser Gly  
305 310 315 320

Ser Asn Phe Cys Ser Glu Gln Lys Thr Ser Gly Ile Ile Asn Lys Phe  
325 330 335

Cys Ser Ala Lys Asp Ser Glu His Asn Glu Lys Tyr Glu Asp Thr Phe  
340 345 350

Leu Glu Ser Glu Glu Ile Gly Thr Lys Val Glu Val Val Glu Arg Lys  
355 360 365

Glu His Leu His Thr Asp Ile Leu Lys Arg Gly Ser Glu Met Asp Asn  
370 375 380

Asn Cys Ser Pro Thr Arg Lys Asp Phe Thr Gly Glu Lys Ile Phe Gln  
385 390 395 400

Glu Asp Thr Ile Pro Arg Thr Gln Ile Glu Arg Arg Lys Thr Ser Leu  
405 410 415

Tyr Phe Ser Ser Lys Tyr Asn Lys Glu Ala Leu Ser Pro Pro Arg Arg  
420 425 430

Lys Ala Phe Lys Lys Trp Thr Pro Pro Arg Ser Pro Phe Asn Leu Val  
435 440 445

Gln Glu Thr Leu Phe His Asp Pro Trp Lys Leu Leu Ile Ala Thr Ile  
450 455 460

Phe Leu Asn Arg Thr Ser Gly Lys Met Ala Ile Pro Val Leu Trp Lys  
465 470 475 480

Phe Leu Glu Lys Tyr Pro Ser Ala Glu Val Ala Arg Thr Ala Asp Trp  
485 490 495

Arg Asp Val Ser Glu Leu Leu Lys Pro Leu Gly Leu Tyr Asp Leu Arg  
500 505 510

Ala Lys Thr Ile Val Lys Phe Ser Asp Glu Tyr Leu Thr Lys Gln Trp  
515 520 525

Lys Tyr Pro Ile Glu Leu His Gly Ile Gly Lys Tyr Gly Asn Asp Ser  
530 535 540

Tyr Arg Ile Phe Cys Val Asn Glu Trp Lys Gln Val His Pro Glu Asp  
545 550 555 560

His Lys Leu Asn Lys Tyr His Asp Trp Leu Trp Glu Asn His Glu Lys  
565 570 575

Leu Ser Leu Ser  
580

<210> 39  
<211> 294  
<212> PRT  
<213> Deinococcus radiodurans

<400> 39  
Gln Leu Gly Leu Val Cys Leu Thr Val Gly Pro Glu Val Arg Phe Arg  
1 5 10 15

Thr Val Thr Leu Ser Arg Tyr Arg Ala Leu Ser Pro Ala Glu Arg Glu  
20 25 30

Ala Lys Leu Leu Asp Leu Tyr Ser Ser Asn Ile Lys Thr Leu Arg Gly  
35 40 45

Ala Ala Asp Tyr Cys Ala Ala His Asp Ile Arg Leu Tyr Arg Leu Ser  
50 55 60

5022704  
Ser Ser Leu Phe Pro Met Leu Asp Leu Ala Gly Asp Asp Thr Gly Ala  
65 70 75 80  
Ala Val Leu Thr His Leu Ala Pro Gln Leu Leu Glu Ala Gly His Ala  
85 90 95  
Phe Thr Asp Ala Gly Val Arg Leu Leu Met His Pro Glu Gln Phe Ile  
100 105 110  
Val Leu Asn Ser Asp Arg Pro Glu Val Arg Glu Ser Ser Val Arg Ala  
115 120 125  
Met Ser Ala His Ala Arg Val Met Asp Gly Leu Gly Leu Ala Arg Thr  
130 135 140  
Pro Trp Asn Leu Leu Leu His Gly Gly Lys Gly Arg Gly Ala  
145 150 155 160  
Glu Leu Ala Ala Leu Ile Pro Asp Leu Pro Asp Pro Val Arg Leu Arg  
165 170 175  
Leu Gly Leu Glu Asn Asp Glu Arg Ala Tyr Ser Pro Ala Glu Leu Leu  
180 185 190  
Pro Ile Cys Glu Ala Thr Gly Thr Pro Leu Val Phe Asp Ala His His  
195 200 205  
His Val Val His Asp Lys Leu Pro Asp Gln Glu Asp Pro Ser Val Arg  
210 215 220  
Glu Trp Val Leu Arg Ala Arg Ala Thr Trp Gln Pro Pro Glu Trp Gln  
225 230 235 240  
Val Val His Leu Ser Asn Gly Ile Glu Gly Pro Gln Asp Arg Arg His  
245 250 255  
Ser His Leu Ile Ala Asp Phe Pro Ser Ala Tyr Ala Asp Val Pro Gln  
260 265 270  
Ile Glu Val Glu Ala Lys Gly Lys Glu Glu Ala Ile Ala Ala Leu Arg  
275 280 285  
Leu Met Ala Pro Phe Lys  
290

<210> 40  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 40  
cacagactcc ctctgtcata ggtttgagtt tatatggaa 39

<210> 41  
<211> 39  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 41  
ttccatataaa actcaaacct atgacagagg gagtctgtg 39

<210> 42  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 42  
cacagactcc ctctgtcata ggttcatgag tttatatgga a 41

<210> 43  
<211> 43  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 43  
cacagactcc ctctgtcata ggttcacatg agtttatatg gaa 43

<210> 44  
<211> 45  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

062206-052704

<400> 44  
cacagactcc ctctgtcata gttcacaca tgagttata tggaa 45

<210> 45  
<211> 47  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 45  
cacagactcc ctctgtcata gttcacaca catgagtttata tatggaa 47

<210> 46  
<211> 53  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 46  
cacagactcc ctctgtcata gtttgagtac tagtactctg agtttatatg gaa 53

<210> 47  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 47  
cgtagaact ccgtcacgaa ttaagcaatt agtaatgcat t 41

<210> 48  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 48  
aatgcattac taattgctta attcgtgacg gagttctaac g 41

<210> 49  
<211> 43  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
oligonucleotide

<400> 49  
cgtagaact ccgtcacgaa ttaagcaatt caagtaatgc att 43

<210> 50  
<211> 45  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 50  
cgtagaact ccgtcacgaa ttaagcaatt cacaagtaat gcatt 45

<210> 51  
<211> 47  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 51  
cgtagaact ccgtcacgaa ttaagcaatt cacacaagta atgcatt 47

<210> 52  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 52  
cgtagaact ccgtcacgaa ttaagcaatt cacacacaag taatgcatt 49

<210> 53  
<211> 41  
<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 53

cgttacaagt ccgtcacgaa ttaagcaatt cgtaacgcat t

41

<210> 54

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 54

aatgcattac gaattgctta attcgtgacg gacttgtaac g

41

<210> 55

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 55

cgttacaagt ccgtcacgaa ttaagcaatt cgtaacgcat t

41

<210> 56

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 56

aatgcgttac aaattgctta attcgtgacg gacttgtaac g

41

<210> 57

<211> 41

<212> DNA

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 57  
cgttacaagt ccgtcacgaa ttaagcaatt cgtaacgcat t

41

<210> 58  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 58  
aatgcgttac gaattactta attcgtgacg gacttggtaac g

41

<210> 59  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 59  
cgttacaagt ccgtcacgac ttaagcaatt cgtaacgcat t

41

<210> 60  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 60  
aatgcgttac gaattgctta aatcgtgacg gacttggtaac g

41

<210> 61  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 61  
cgttacaagt ccgtcacgaa ttaagcaatt cgtaacgcat t 41

<210> 62  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 62  
aatgcgttac gaattgctta attcgtaacg gacttgtaac g 41

<210> 63  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 63  
cgttacaagc ccgtcacgaa ttaagcaatt cgtaacgcat t 41

<210> 64  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 64  
aatgcgttac gaattgctta attcggtacg gacttgtaac g 41

<210> 65  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 65  
cgttccaagt ccgtcacgaa ttaagcaatt cgtaacgcat t 41

<210> 66  
<211> 41  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 66  
aatgcgttac gaattgctta attcgtgacg gacttgaaac g

41

<210> 67  
<211> 78  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
oligonucleotide

<400> 67  
tgcacggacc tcgagagaac cgaacgatcc ggtatgtatc atgggacaga attcgacac 60  
tgcactcgac gatagtct 78

<210> 68  
<211> 49  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 68  
tgcagtgtgc gaattctgtc ccatcatcac atccggatcg ttccgttct

49

<210> 69  
<211> 62  
<212> DNA  
<213> Artificial Sequence

<220>  
<223> Description of Artificial Sequence:  
Oligonucleotide

<400> 69  
gactatcgtg cagtgcagtg tgcgaattct gtcccatcat cacatccgga tcgttcggtt 60  
ct 62

<210> 70

```
<211> 64
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide

<400> 70
tgcagtgtgc gaattctgtc ccatcatcac atccggatcg ttcgggttctc tcgaggtccg 60
tgca                                         64

<210> 71
<211> 79
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
      Oligonucleotide

<400> 71
gagactatcg tgcagtgcag tgtgcgaatt ctgtcccatc atcacatccg gatcggttcgg 60
ttctctcgag gtccgtgca                                         79
```